Timbre as Vertical Process: Attempting a Perceptually Informed Functionality of Timbre

Anthony Tan

McGill University, Department of Music Research (Composition)
Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT)

anthony.tan@mail.music.mcgill.ca

Abstract

This paper suggests a perceptually informed compositional model of timbre functionality. A functionality of timbre may refer to the compositional control of timbre in order to generate experiences of *tension*, the perception of which, serves as an important link between the recognition of *form-bearing dimensions* and subjective emotional response. Utilizing the psychoacoustic models of *Spectral Fusion* (McAdams, 1982) and *Auditory Scene Analysis* (Bregman, 1994), parameters that affect the formation or segregation of a perceptually unified timbre source are compositionally manipulated. Two attempts at the model are suggested and discussed. The paper ends with a discussion of the micro-temporal control of timbre.

Context

Within electroacoustic music studies, numerous methods to identify, describe and represent timbre structures have been explored. Less explored, however, is an effective codified model to use timbre *functionally* as a *form-bearing dimension*[1]. Part of the difficulty in achieving this task is that timbre represents a multidimensional parameter, both in its definition and description. In addition, timbre maintains a paradoxical function in daily human life. Timbre arises from a distributed perception, as we lack a dedicated physiological mechanism for its perception (Fales, 2002). As such, the acoustic stimulus and the perceived stimulus are incongruent. Paradoxically, we use this perceptual product to define our acoustical world. One might say that, from a universal sense, object recognition represents timbre's primary role. Musically, the shift from considering timbre as a carrier of pitch and rhythmic information, towards timbre as the primary object of musical meaning, represents one of the primary facets of electroacoustic music (Erickson, 1975). The perceptual nature of timbre, however, poses challenges when, as composers, we aim to achieve strict parametrical control of timbre, as the cognitive mechanism constrains the resulting auditory image. Therefore, a functional model of timbre developed from the parametrical control of timbre parameters must arise from a perceptual point of view.

[1] Form bearing dimensions are perceptually differentiated musical forms as opposed to purely structural constructions employed by the composer.

Structure versus Function

Pierre Schaeffer (Schaeffer *et al.*, 1967) may be considered the first musical thinker to break away from connecting timbre to its source and define objective timbre structures. The *Objet Sonore* gave composers a method to characterize perceived timbre units of a specific temporal scale and describe these units according to their perceived qualities. This paradigm proved to be fruitful as the idea of *perceptual object* pervades much electroacoustic music discourse. We can also define a timbre structure based on multidimensional timbre scaling (Grey, 1977). Here, dissimilarity judgments between spectral frequency and spectral temporal qualities position a timbre within a multi-dimensional space.

These methods of identifying timbral structures reveal, however, very little as to how we may use these structures in a functional manner. By definition, a functional structure should serve as a 'vehicle for expressivity' (Berry, 1976). Thus, within electroacoustic music, how a timbre structure (sound-object) operates within a context or a paradigm in order to generate varying experiences of tension becomes increasingly important.

Tension versus Dissonance

Tension may be defined as a perceptual quality that acts as a link between the perception of a musical structure (sound object) as a form bearing dimension and subjective emotional response. Increasing tension is often qualitatively described as increasing excitement, expectation, or uncertainty, while the decrease in tension may be described as relaxation, fulfillment, or resolution. Equally important is the separation of *dissonance* from *tension*. As the primary aspect of tension, *Dissonance,* may further divide into two types: *Sensory Dissonance*, or *Syntactic Dissonance*. With sensory dissonance, the sensation of 'roughness' occurs when two tones with an interval smaller than the critical bandwidth, sounded simultaneously create 'beats'. There exists a strong relationship between the sensation of beating and the perception of dissonance (Helmholtz ,1885; Bigand, Parncutt and Lerdahl, 1996).

One may also consider a syntactic approach to dissonance whereby defined rules and processes engender *dissonance* from a context of relationships. A general definition of syntactic dissonance could be considered as the movement between "stability and instability" (Cazden, 1980). Within electroacoustic music one often discusses the movement between recognizability of sound objects to non-recognizability, or even gestural expectation and suspension to that of completion and resolution.

Historical Approaches

In general we may divide approaches to timbre functionality in three broad categories: linguistic/semiotic, gestural, or as an extension of harmony. These categories, however, are not mutually exclusive and many models of timbre functionality involve combinations of these categories. Furthermore, the authors of these proposed models do not necessarily use the term *Timbre Function*, but do address the use of timbre as a dimension able to produce varying experiences of tension.

A linguistic/semiotic approach either uses models from linguistics applied to music, or though the creation of symbolic representations of timbre structures, a meta-language in order discuss timbre functions emerges. In *Timbre Hierarchies* (Lerdahl, 1987), *Linguistic Parse Trees* deploy *Timbre Prolongation Structures* in which a sense of belongingness or separateness arises from a comparison to a timbral prototype. Wessel (1979), discusses the control of a multidimensional timbre space in order to generate timbre structures analogous to pitch structures. Semiotic approaches derive from Jean-Jacques Nattiez's analyses at the *Neutral Level* to create identifiable perceptual units in a work (Nattiez, 1990). Using this, Stephane Roy in his *L'analyse des musiques électroacoustiques, modèles et propositions* (Roy, 2004) defines a *Functional Grid* of four main categories (orientation, stratification, process, rhetoric) sub-classified into symbols of forty-five functions.

From the point of view of gesture, tension functions are defined in terms of teleological energy profiles over time. *Spectromorphology* (Smalley, 1997), which deals with the shaping of spectral components, represents a prevalent gestural approach to timbre function. Smalley identifies various motion and growth processes that denote different forms of spectral movement over time relating to expectation and resolution. Smalley also defines a form of tension arising from the cognitive dissonance created by the perceived link between the source and cause of a timbre, which he defines as *gestural surrogacy*.

Another gestural approach employs concepts from embodied music cognition. Here, the body serves as a mediator between heard musical structures and the cognitive patterns that arise in the mind. Godøy defines *Gestural Sonorous Objects* (Godøy, 2006) whereby the perception of sound objects is linked to a process of re-enactment by the physical body. Therefore timbre function could be considered a result of direct embodied experiences of gestural expectation and resolution. Furthermore, the *Unités Sémiotiques Temporelles* (Delalande *et al.*, 1996) combines both semiotic and gestural approaches, creating a bio-semiotic analysis such that timbre structures are identified and nominalized based on physical energy profiles.

Finally an approach to timbre function may be seen as an extension of harmony. Kaija Saariaho defines a sound-noise axis in which harmonic sounds are considered consonant and noise/inharmonic sounds are considered dissonant (Saariaho, 1987). The harmonic approach further brings in aspects of sensory dissonance. Parncutt and Strasburger (1994), as well as Shields and Kendall (2004), have discussed the measurement of sensory dissonance on harmonic and non-harmonic structures as a means of controlling tension within a musical work.

Theoretical Background

As timbre arises from perceptual processes, perceptual aspects should also determine timbre functionality. The functional approach this paper proposes utilizes sensory dissonance models in relationship to two other psychoacoustic models. Firstly, *Spectral Fusion* (McAdams, 1982), denotes the process by which spectral information fuses together to form one single perceptual object that we define as being a single heard timbre. Further, McAdams distinguishes between two types of listening: *synthetic* or *analytic listening*. Perceiving one fused timbral object is defined as *synthetic listening* while perceiving independent spectral components is termed *analytic listening*.

Auditory Scene Analysis (Bregman, 1994) represents the second perceptual theory the proposed model employs. Auditory stream formation theory deals with how the auditory system determines whether incoming acoustic information results from one, or more than one, 'source'. In our case, this 'one source' is a spectrally fused sound object. Understanding the manner in which composers manipulate the streaming process in order to create synthetic or analytic listening situations represents one of the primary goals of this paper. "Composers can manipulate the pattern context in which spectral structures appear, according to the set of principles which govern the fusion and segregation of partials by the auditory system." (Wright, Bregman, 1987).

It has been shown that, from a tonal perspective, timbre differences affect listeners' perception of sensory dissonance and fusion (Shields, Kendall, 2004). The relationship, however, between the fusion and segregation of *spectral* components to the perceived sensory dissonance of a timbre structure remains less clear. In an experiment utilizing Webern's orchestration of Bach's *Ricercar*, researchers hypothesized a relationship between sensory dissonance and auditory streaming (McAdams, Paraskeva, 1997). Using two forms of the piece, an orchestrated version and a direct piano transcription of the orchestrated version, respondents in this study continually rated the piano version as having a higher degree of perceived tension. The hypothesis was that the lesser perceived dissonance of the orchestral version results from timbre-induced stream segregation. *Auditory scene analysis* processes group acoustic information into streams after which auditory roughness is computed (McAdams, 2012). With the single piano timbre, which represents one fused spectral object, *sensory dissonance* computation occurs on the frequency components of a single auditory stream. In other words, due to the spectrally fused piano timbre, dissonance arising from the intervallic content of the music was calculated as belonging to one single source. With the orchestral version, timbre differences among orchestral instruments dissipate sensory dissonance among multiple streams. This suggests the role of timbre as a form-bearing dimension, able to produce differences in perceived tension.

Compositional Rule?

One might propose that the fusion of complex timbres, or the spectral components of a single timbre, into a single blended auditory image generates a higher degree of sensory dissonance than if the complex timbres, or spectral components of a single timbre, segregate into separate auditory streams. In order to explore this hypothesis, we must first define the parameters that affect spectral fusion or segregation. Synthesizing the work of (McAdams, 1982; Bregman, 1994; Shields, Kendall, 2004), the proposed parameters that affect spectral fusion are:

- Harmonicity of spectral components (phase locking);
- Congruency of modulation between spectral components;
- Attack/onset synchronicity of spectral components (temporal and spatial);
- Similarity of spectral centroid;
- Familiarity of formant structure.

These fusion parameters represent controllable compositional parameters that affect the fusion or segregation of spectral components. The next step would be to plot these parameters, linearly, on a continuum:

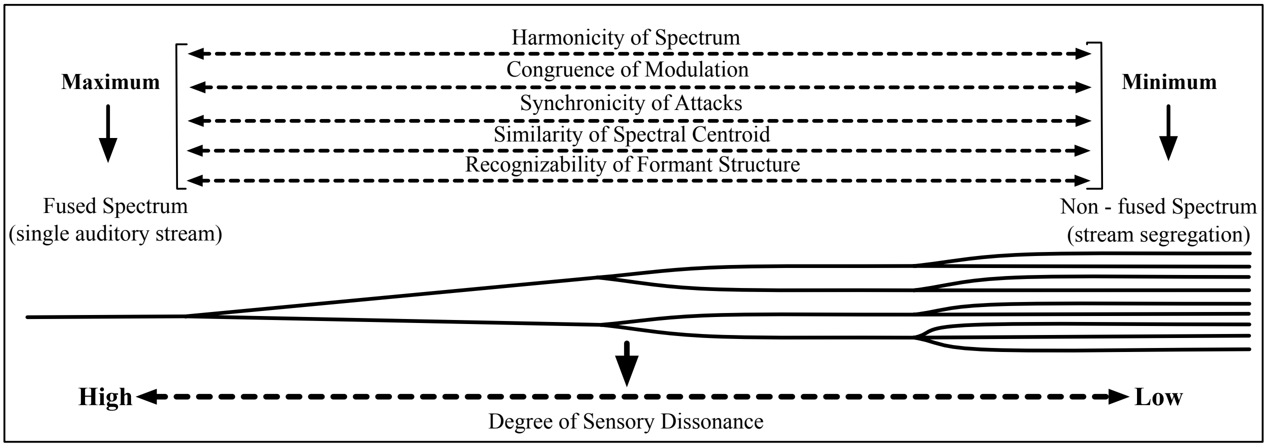

Figure 1: Continuum of Fusion Parameters

On the left, when one has a maximum of harmonicity among spectral components, a maximum congruence of modulation, the maximum synchronicity of attack, the maximum similarity in spectral centroid, and a formant structure that is extremely recognizable, this leads to an optimal situation of spectral fusion and perhaps a higher degree of sensory dissonance. Conversely, with the minimum amount of these parameters, we would have a situation of a non-fused spectrum, leading to stream segregation, and therefore perhaps a lower degree of sensory dissonance.

Fusion parameters, however, need not be coupled together and the independence of parameters would generate varying levels of fusion. In order to approach this complexity one may adapt from Lerdahl's *Timbre Hierarchies* (1987), the idea of multi-dimensional timbre arrays.

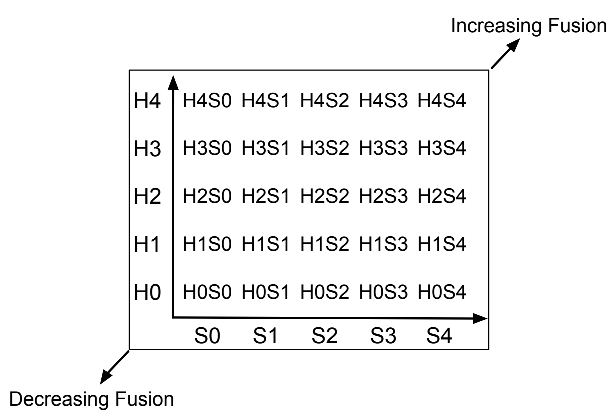

Figure 2: Bi-dimensional timbre array for harmonicity and synchronicity

In figure 2, as an example, by plotting two parameters on an axis, harmonicity and synchronicity, one creates a bi-dimensional timbre matrix whereby each unit represents varying levels of fusion. Each step in one dimension of an array represents what psychophysicists define as a *just* *noticeable difference* in that particular parameter. On the Y-axis, moving from H0 to H4 represents an increase in harmonicity. On the X-axis, moving from S0 to S4 represents an increase in synchronicity of attack transients of spectral components. Within the matrix, the top right unit, H4S4, represents a timbre structure with the highest amount of harmonicity and synchronicity and therefore the highest amount of fusion. Conversely, H0S0, at the bottom left represents the least fused timbral structure and perhaps the least sensory dissonant. Next, by moving linearly along the diagonal axis, we can extract a fusion array:

F0 → F1 → F2 → F3 → F4

H0S0 H1S1 H2S2 H3S3 H4S4

Figure 3: Fusion Array

A fusion array could be considered a higher order array that acts as an emergent property from the manipulation of lower level parameters. Intervals of fusion present a scale of fusion units moving from least to highest level of fusion and perhaps from a lower to higher level of sensory dissonance. Also important is the selection of a psychological prototype. Many perceptual categories contain units that can be considered the most central, or most stable, or the most representative unit of that category (Lerdahl, 1987). For our purposes the selection of the most fused member of our fusion array (F4) would be considered the fusion prototype serving as a psychological anchor.

Attempt I

In an initial attempt, I began with a single complex timbre of my own construction. This complex timbre had a fundamental frequency of seventy-five Hz (or 'D') and a temporal length of approximately eleven seconds. It was important to define a temporally 'closed' sound object as my timbral prototype. This timbral prototype also represented, from my perception, a perceptually fused object. Minor noise elements were also added to the sound object to represent the *specificities* of instrumental sounds. *Specificities* are unique timbral elements of instruments such as the click of the keys on a wind instrument, or the sound of the plectrum plucking the string on a harpsichord.

By using spectral filtering software I separated the spectral components up to the 16th partial while leaving the remaining upper partials, past the 16th partial, as a separate stream. I aimed to manipulate parameters of synchronicity and harmonicity in order to control fusion. The fusion array, (F4 → F0 → F1 → F2 → F3→ F4), begins with the most fused/tense unit dropping down to the least and then building up incrementally to the most fused/tense unit of the array. In Figure 4, one can observe the direct relationship between fusion levels and the number of intended auditory streams. By separating different portions of the spectrum, groups of the spectrum can temporally shift, controlling synchronicity. Frequency transposition (in semi-tones) of the various streams was used to control harmonicity with the fundamental and between spectral components.

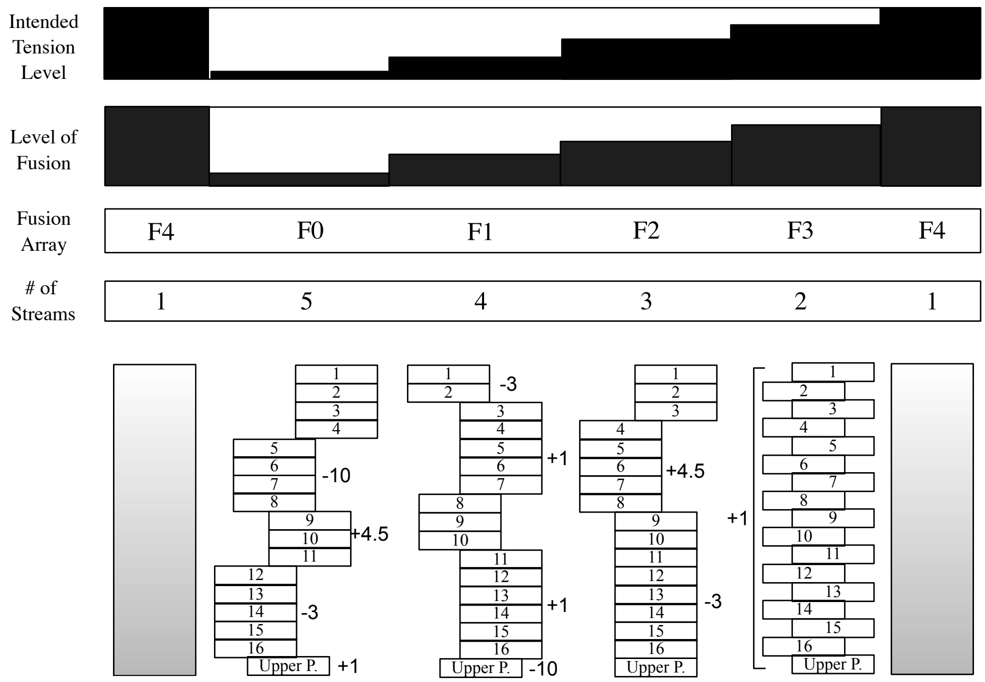

Figure 4: Attempt I

Discussion of Attempt I

Various issues arose during this attempt. Firstly, the temporal distance between spectral groups (affecting synchronicity) was too large leading to sequencing effects. As such, one hears melodic contours instead of shifting spectral groups. This points to the importance of distinguishing between a pitch listening situation versus a timbre listening situation (Erickson, 1975). Another problem was related to the *specificities*, or noise elements added to the sound object. *Specificities* do not combine easily with harmonic components. Thus an understanding of how noise and pitch elements relate to fusion need to be further explored. *Specificity* segregation became more evident when other, more prominent, spectral components were no longer synchronized. Finally, and perhaps most importantly, context effects became apparent in this situation. Whether one sets up a *synthetic* or an *analytic listening* situation will determine how one perceives subsequent structures. As the spectrum becomes separated through the course of this attempt, it leads the ear to an *analytic listening* situation. Subsequently at the end, when the original fused object is presented, we remain in this *analytic listening* situation, perceiving separate spectral components of the previously heard fused sound object.

Attempt II

The second attempt aimed to fuse multiple complex timbres, into an emergent complex timbre. Each individual complex timbre was a short impulse of seventy milliseconds. A fundamental frequency of 183 Hz (or 'F♯') was chosen and all the sound files were transposed such that their fundamental frequencies coincided. The individual complex timbres were single attacks on the following instruments:

- A gong placed on top of a timpani;
- A piano cluster in the mid range;
- A struck large metal plate with a medium hard mallet;
- An electric guitar with distortion;
- Wooden drum sticks struck against each other.

By controlling the amplitude and the fundamental frequency of each complex timbre, a fusion prototype was created. Using a constant repetition of the fusion prototype at 130 beats per minute, signal processing algorithms were then employed on each individual instrument in order to affect fusion dynamically over time. *Tremolo* controlled the congruency of amplitude modulation. A *pitch shifter* affected the harmonicity of the individual timbres within the timbre complex. A *delay* line controlled the synchronicity of attack. A *band bass filter* altered the frequency distribution of the individual timbres, thereby controlling spectral centroid. The electric guitar and the wooden drumsticks were kept constant throughout, serving as an acoustic 'glue'.

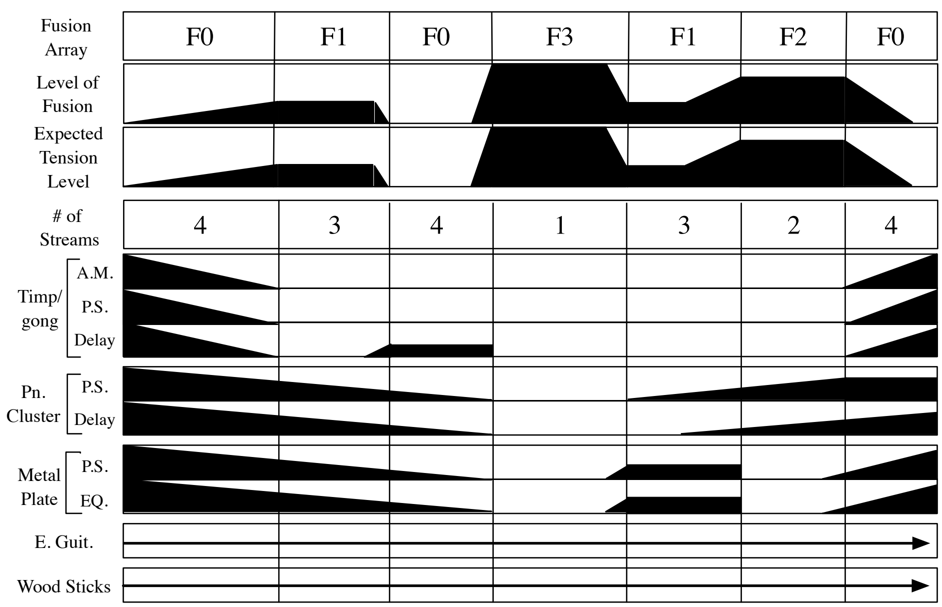

Figure 5: Attempt II

Discussion of Attempt II

In this second attempt a question that arose is to what degree could the individual complex timbres be heard as separate elements but still maintain the perception of a complex sound object? In other words, when do we perceive the parts and no longer the whole? This raises the connection between perceptual fusion and the Gestalt concept of 'belongingness'. There exists a perceptual threshold in which altering the individual timbres with signal processing algorithms segregated the timbre complex into individual timbres but still maintained the perception that the separate timbres belonged together. Moving past this perceptual threshold engenders all individual timbres to be independently heard with no complex structure perceived.

Nonetheless, one can say that signal-processing algorithms directly affected the fusion of the timbre prototype and the independent variation of these variables alters the number of perceived auditory streams. Synchronicity of attack seemed to be the most effective. Delay times over 90 ms, however, led to more rhythmic effects such that individual complex timbres began to emerge, moving us away from a timbre listening situation to a rhythmic listening situation. Pitch shifting beyond 200 cents in either direction began to create melodic effects, leading to a pitch listening situation. Altering the frequency distribution with a *high shelf filter* seemed to be most affective if the higher portion of the spectrum above 3 000 Hz was attenuated to a large degree, affecting 'brightness'.

General Discussion

Obviously my attempts were preliminary examples of what I think could be a very refined model of functionality. The compositional control of fusion parameters definitely altered *synthetic* or *analytic listening* situations. It is difficult to say, however, whether or not the fused timbre prototype could be perceived as more dissonant than a segregated spectrum. Subsequent studies should clearly measure the sensory dissonance over time in order to accurately predict the model. One fusion parameter not fully explored was the recognizability of the formant structure. Recognized formant structures are fused timbre structures that we identify from a learned catalogue of timbre structures. Perhaps this relates to Denis Smalley's idea of *source bonding* (Smalley, 1994), whereby a higher recognizability of formant structure leads to a higher degree of source bonding. Non-recognized formant structures could lead the cognitive mechanism to stream segregation and non-recognizability. We also need to define which parameters of fusion are more salient. For example, does synchronicity have more effect on fusion than modulation congruency? As well, the level and manner of interaction between these parameters could have effects on the perceptual fusion of a sound object. The threshold of signal processing effects on fusion needs to be further clarified.

Furthermore, we need to begin to understand the interaction between pitch and timbre. Pitch is much more salient, especially in western culture. Too much pitch variance would draw the perception away from timbre listening. At the same time, this model requires an understanding of frequency intervals and the sensory dissonance that arises from their interaction. As well, context effects need to be clearly understood. The order of presentation of spectral components affects whether or not we use *synthetic* or *analytical listening* strategies. Beginning with *analytical listening* modes tends to keep that listening mode in place. Finally, my fusion attempts were very linear. Most composers play with varying levels of fusion among timbre structures within a piece. It would be interesting to develop *Fusion hierarchies* in order to create a rich and complex functional model.

Timbre as Vertical Process

This work suggests that an aspect of timbre's functionality is in the vertical dimension. Meaning, timbre is a process that happens *in* time versus *over* time. Di Scipio (1994) previously discusses this in his idea of *micro-time sonic design*. Although we may track differences over time, the emergent property that we call timbre occurs within a micro temporal interval. In this way, one may view timbre as being a vertical process in music. This vertical process represents the parametrical control of fusion parameters such as attack synchronicity, and harmonicity, within a micro-temporal time frame. Experimental evidence for this verticality exists. In an experiment addressing the relationship between timbre and pitch, Krumhansl and Iverson observed that timbre is perceived as a situational parameter not relational (Krumhansl, Iverson, 1992). Meaning, timbre perception is not intervallic as analogous to pitch structures but rather, absolute and in the moment. This may be referred to as the competition between horizontal and vertical forces. "Horizontal forces create the sequential patterns that we hear – pitch order and rhythm. It is the vertical force that creates perceived qualities such as timbre and consonance and dissonance that emerge from the fusion of simultaneous components into a single sound." (Wright, Bregman, 1987)

Finally, it should be clearly stated that fusion is not another parameter of timbre. Fusion *is* timbre. The horizontal/sequential and vertical/simultaneous processes of perceptual organization, work together to form a fused auditory stream. Timbre is the perceptual description of this fused auditory stream (Bregman, Pinker, 1978).

Conclusion – Implications

The fusion or segregation of spectral components can be compositionally controlled through specific fusion parameters. The compositional control of the fusion process has effects on the overall perceived sensory dissonance of a sound object. Thus, spectral fusion could be considered an aspect of timbre functionality. There are certain implications with this proposition. Firstly, this proposition does not eradicate the other functional models, such as the gestural or semiotic approaches. As timbre is multi-dimensional, therefore its functionality is also multi-dimensional. The fusion or non-fusion of spectral properties merely presents another dimension to this complex situation. Also, the concept of fusion adds another category to the *objet sonore*. The *objet sonore* is really an *objet perceptive,* and whether or not that object is perceived as being a fused spectral structure or segregated one, will have effects on the overall dissonance of that object. Next, signal processing has a direct effect on fusion parameters. How one uses signal processing and its effect on the fusion of segregation of partials could have effects on the overall sensory dissonance of spectral structures. Finally, perhaps we are just at the beginning of the development of a timbre counterpoint. The connection between perception and musical systems has always been present. The rules of tonal counterpoint also developed from the avoidance of dissonant intervals that have sensory dissonance correlations. Perhaps we are now beginning to develop our own contrapuntal rules in regard to timbre.

References

BERRY Wallace, *Structural Functions in Music*, Englewood Cliffs (NJ, USA), Prentice-Hall, 1976.

BIGAND Emmanuel, PARNCUTT Richard and Fred LERDAHL, “Perception of Musical Tension in Short Chord Sequences: The Influence of Harmonic Function, Sensory Dissonance, Horizontal Motion, and Musical Training”, *Perception & Psychophysics*, 58, 1996, pp. 125-141.

BREGMAN Albert S., *Auditory Scene Analysis: The Perceptual Organization of Sound*, Cambridge (MA, USA), The MIT press, 1994.

BREGMAN Albert S., PINKER Steven, “Auditory Streaming and the Building of Timbre”, *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 32(1) 1978, pp. 19-31.

CAZDEN Norman, “The Definition of Consonance and Dissonance”, *International Review of the Aesthetics and Sociology of Music* 11(2), 1980, pp. 123-168.

DELALANDE François *et al.*, *Les Unités Sémiotiques Temporelles*, Laboratoire Musique et Informatique de Marseille (MIM), Documents Musurgia, Paris, Édition Eska, 1996.

DI SCIPIO Agostino, “Micro-Time Sonic Design and Timbre Formation”, *Contemporary Music Review*, 10(2), 1994, pp. 135-148.

ERICKSON Robert, *Sound Structure in Music*, Berkeley and Los Angeles (CA, USA), University of California Press, 1975.

FALES Cornelia, “The Paradox of Timbre”, *Ethnomusicology*, 46(1), 2002, pp. 56-95.

GODØY, Rolf Inge, “Gestural-Sonorous Objects: Embodied Extensions of Schaeffer’s Conceptual Apparatus”, *Organised Sound*, 11(02), 2006, pp. 149-157.

GREY John M., “Multidimensional Perceptual Scaling of Musical Timbres”, *The Journal of the Acoustical Society of America*, 61(5), 1977, pp. 1270-1277.

HELMHOLTZ, Hermann von, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, translated by Alexander J. Ellis, London, Longmans, Green and Co, 1885.

KRUMHANSL Carol L., IVERSON Paul, “Perceptual Interactions between Musical Pitch and Timbre”, *Journal of Experimental Psychology: Human Perception and Performance*, 18(3), 1992, pp. 739-751.

LERDAHL Fred, “Timbral Hierarchies”, *Contemporary Music Review*, 2(1), 1987, pp. 135-60.

MCADAMS Stephen, “Spectral Fusion and the Creation of Auditory Images”, in *Music, Mind, and Brain*, Manfred Clynes (ed.),1982, pp. 279-298.

MCADAMS Stephen, “Musical Timbre Perception”, in *The Psychology of Music*, Diana Deutsch (ed), 2013, 3rd edition, pp. 35-67.

MCADAMS Stephen, PARASKEVA Stella,. "Influence of Timbre, Presence/absence of Tonal Hierarchy and Musical Training on the Perception of Musical Tension and Relaxation Schemas", In *Proceedings of the 1997 International Computer Music Conference, 25-30 September 1997*, *Thessaloniki, Hellas,* 1997, pp. 438-441.

NATTIEZ Jean-Jacques, *Music and Discourse: Toward a Semiology of Music*, translated by Carolyn Abbate, Princeton (NJ, USA), Princeton University Press, 1990.

PARNCUTT Richard, STRASBURGER Hans, "Applying Psychoacoustics in Composition: 'Harmonic' Progressions of 'Nonharmonic' Sonorities", *Perspectives of New Music*, 32(2), 1994, pp. 88-129.

ROY Stéphane, *L'analyse des musiques électroacoustiques : Modèles et propositions*, Paris, Éditions L'Harmattan, collection "Univers Musical", 2003.

SAARIAHO Kaija, "Timbre and Harmony: Interpolations of Timbral Structures", *Contemporary Music Review*, 2(1), 1987, pp. 93-133.

SCHAEFFER Pierre, *La Musique Concrète*, Paris, Presses Universitaires de France, collection "Que sais-je", n° 1287, 1967.

SHIELDS Robin, KENDALL Roger, "The Relation of Timbre to Dissonance and Spectral Fusion", in *Proceedings of the 8th International Conference on Music Perception & Cognition*, Evanston (IL, USA), 2004, pp. 596-599.

SMALLEY Denis, "Defining Timbre – Refining Timbre", *Contemporary Music Review*, 10(2), 1994, pp. 35-48.

SMALLEY Denis, "Spectromorphology: Explaining Sound-Shapes", *Organised Sound*, 2(02), 1997, pp. 107-26.

WESSEL David L., "Timbre Space as a Musical Control Structure", *Computer Music Journal*, 3(2), 1979, pp. 45-52.

WRIGHT James K., BREGMAN Albert S., "Auditory Stream Segregation and the Control of Dissonance in Polyphonic Music", *Contemporary Music Review*, 2(1), 1987, pp. 63-92.